DOCUMENT RESUME

ED 072 472 CS 500 119

AUTHOR Gruber, Frederic A.
TITLE Auditory and Articulatory Aspects of Grammar.
PUB DATE Dec 72
NOTE 15p.; Paper presented at the Annual Meeting of the Speech Communication Assn. (58th, Chicago, December 27-30, 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Acoustic Phonetics; *Auditory Perception; Language Research; *Linguistics; Linguistic Theory; *Pattern Recognition; Phonetic Analysis; *Speech; Transformation Generative Grammar

ABSTRACT

The author addresses the need for a new acoustic recognition strategy, extending the position that any adequate grammar of a language must distinguish between auditory and articulatory knowledge. Reviewing the existing literature and theories of language and its acquisition, the author discusses their limitations and inadequacies in accounting for natural acoustic patterns. As a solution to the acoustic recognition problem he proposes differentiation of speech signals based on time and intensity. This would provide unique representations for input syllables according to their linguistic differences. (Author/LG)

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 072472

AUDITORY AND ARTICULATORY ASPECTS OF GRAMMAR*

by

Frederic A. Gruber

SUNY-Buffalo

and

University of Minnesota

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Frederic A. Gruber
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

presented before the Speech Sciences Division of the Speech Communication Association at the 58th annual meeting in Chicago, Illinois, December 28, 1972

*This work was, in part, supported under a William Stout Fellowship from the Graduate School, University of Minnesota, Minneapolis and by the facilities of the Wendell Johnson Speech and Hearing Center, The University of Iowa, Iowa City.

CS 500 119

A. Linguistic and Philosophical Background

Generative-transformational grammars proposed by Chomsky (1957, 1965) Katz and Postal (1964) and others, are alleged to be principled expressions of the facts of language underlying any particular utterance as a token of that language. The analogy between the body of mathematical knowledge accumulated by generations of mathematicians and expressions from that body of knowledge in concrete calculations of a given mathematician is used pedagogically to illustrate the nature of the competence/performance dichotomy. However, one should not confuse the mathematics analogy with the linguistic distinction. Any mathematical formualtion is capable of paraphrase in natural language; yet, we cannot expand studies of generative power to mathematical knowledge as a subset of natural language. The two systems are distinct and different.

We expect each normal child to master its native language in the normal developmental process. We do not similarly expect each child to become naturally "fluent" in higher mathematics. The notion of explanatory adequacy in linguistic theory is deeper than its hypthetical counterpart might be in mathematical theory.

Chomsky describes a grammar of a language by a set of rules which expresses correspondence between sound and meaning in that language. The sense in which "grammar" is applied, however, is alleged to be "loose"; not all aspects of sound and meaning in the ordinary sense of these terms are suitable in a theory of competence (Chomsky, 1971). A more precise expression of the sound-meaning consequences of syntactically motivated grammars terminates with abstract, phonetic representations and semantically significant features. (Chomsky, 1963; 326-330) Such qualification is needed in the absence of resolution of the mind-body problem which has so long perplexed philosophers.

Empirical consequences of generative grammars can, at best, entail only inter-theory compatibility; we would like our grammars to be in substantial agreement with theories of neurology (see Whitaker, 1969), anatomy and acoustics (see Jakobson, et. al., 1969), etc. Resolution of grammatical strings to direct, empirical consequences implies disentanglement of the philosophical weltknoten, rending unnecessary a competence-performance dualism.

As in mathematics, purely descriptive accounts need not be ontologically interesting. However, explanatory attempts as suggested by Chomsky (1965) to be a goal of linguistic theory cannot altogether avoid ontological commitment. What constitutes "natural classes," "similar processes," and, in short, linguistically significant generalizations must be determined (Chomsky, 1965; p. 42).

Chomsky's (1971) expression of "standard" generative transformational theory proposes the generation of quadruples (P,s,d,S) where P is a phonetic representation, s is a surface structure, d is a deep structure, and S is a semantic representation. No ordering of these structures is intended. To seriously suggest such a "direction of mapping" is, as Katz and Postal (1964) remark, a competence-performance category mistake of the most misleading sort. Yet, within each component of the grammar ordered rules are said to be necessary in order to achieve descriptive adequacy. Syntactic structures are usually defined in terms of predicates such as "precedes," "dominates," and "is labeled," so that transformations act as well-formedness constraints on successive phrase markers.

B. The Importance of Global Constraints

In well known arguments, Lakoff (1971) denies the autonomy of the syntactic component with respect to semantics. In Chomsky (1965) and in Chomsky (1971)

all lexical insertion in a derivation must occur prior to transformations. Lakoff (1971) argues convincingly that certain transformations must occur prior to lexical insertion. In consequence, all principled distinction between syntax and semantics is lost in favor of more "global" constraints operating across entire derivations.

A result of this reasoning is the inclusion into grammars of "performance" considerations such as sentential presupposition (Lakoff, 1971), aptness (Fillmore, 1971), and use (Langendoen, 1971). As Maclay (1971) remarks: "The autonomy of competence may well be the next victim."

The careful distinction between a phonological component and syntax has not been attacked as incisively as the syntax-semantics division. But, there is indication that such arguments are to be forthcoming. Sampson (1970) argued in favor of a level of deep structure in phonology to preserve the principle of symmetery as a simplicity criterion. Fudge (1967) took issue with the handling of systematic phonemic and systematic phonetic levels on the same terms, suggesting the more abstract, phonological (systematic phonemic) level ought be the focus of generative theory. Turner (1970) maintained that a separate grammar for the speaker and for the hearer are ultimately needed to account for a realistic perceptual model.

Extension of arguments for global constraints spanning phonology and syntax have been generally resisted for, as Lakoff (1972) remarks, " ...one must consider the naturalness of the linguistic units used in the coding operation. It is generally conceded that the units used in phonological description should have an independent natural basis in phonetics. Phonological rules are taken as using phonetic features, which are given independent of those phonological rules." (p.77)

Yet, in the same article, Lakoff (1972) says, "Clearly, one gains a deeper

insight into the nature of phonology by avoiding non-natural features such as ±ARBITRARY. If avoiding non-natural features leads one to the conclusion that global rules are necessary in phonology, then that is an interesting fact about the phonological structure of natural languages." (p.79)

C. Global Constraints in Phonology

However, Kisseberth (unpublished ms) provides evidence that global rules occur in Klamath phonology. I believe this will be a general trend in phonology. If global derivational constraints span the syntactic-phonological boundaries, a number of consequences immediately follow. (1) The syntax-phonology boundary becomes unnecessary. (2) The competence-performance distinction as it applies at these levels is seriously altered. (3) The basis for "natural" classes and features which are alleged to have an independent basis in phonetics is opened to serious question. (4) A possibility of separate grammars for speaker and hearer emerges.

A number of considerations which favor global constraints across the syntax-phonology boundary. The flood of alternative proposals for phonology testify that formulation in the manner of Chomsky and Halle (1968) has not proven "deep and satisfying." Important among the reasons for dissatisfaction is the embarrassingly ad hoc character of readjustment rules needed to fit terminal strings from the syntactic component to imput conditions required of the phonological component. (Chomsky and Halle, 1968; 10, 371; Turner, 1970). Turner (1970) points out that a way of avoiding this difficulty is by the introduction of phonological constraints all along the process of a derivation. This is equivalent to a global rule in the sense of Lakoff.

Certain types of puns, rhyme, and alliteration whose appropriateness is determined by phonological considerations, even occasionally at the expense of

complete grammaticality, demands lexical selection and insertion after a derivation has reached phonological levels. Moreover, as any parent will testify, an ability to pun and rhyme appears to be a part of normal language acquisition, though it does occur relatively late. Either a case must be made for phonologically motivated selection restrictions, or we must accept the claim that global constraints operate across entire derivations.

D. Naturalness in Perceptual Consequences of Grammars

Lack of well-motivated principle also pervades establishment of correlates for distinctive features, although somewhat appealing taxonomies of sound and articulatory activity have been related to proposed features (Chomsky and Halle, 1968; Jakobson, et. al., 1969). To say there is no principled relation between phonetic representations and actual sound or articulatory gestures is to restate that no impelling solution to the perplexing problem of pattern recognition has been achieved. This is not particularly damning. However, until the problem is, in principle, solved it is not possible to be precise and explicit as to what can be meant by an "independent and natural basis" for phonetics. Moreover, it is not reasonable to claim current taxonomies are a first approximation unless the same principles are capable of closer approximations. Unfortunately, recent work in speech synthesis demonstrates this cannot be so.

For example, Jakobson, Fant and Halle (1969) distinguish the phonemes /b/, /d/ and /g/ by the features grave/acute and compact/diffuse:

	/b/	/d/	/g/
Grave/acute	+	-	
Compact/diffuse	-	-	+

The acoustic gravity feature is identified by, "...observ/ing/ the second formant in the adjacent vowel, if any: it is lowered in the case of grave consonants, and raised if the consonant is acute...in some cases the position of the third and higher formants may also be affected." (p.30)

The acoustic compactness feature is, for consonants, "...displayed by a predominant formant region, centrally located, as opposed to phonemes in which a non-central region predominates." (p.27)

Work on speech perception and synthesis conducted largely at the Haskins Laboratories (see, Delattre, Liberman, and Cooper, 1955) demonstrates such correlates are, at best, crude guesses about the nature of "natural" acoustic patterns of linguistic import.

Chomsky and Halle (1968) abandon the compact/diffuse and grave/acute features in favor of an alternative scheme (pp. 306 passim). However, they neglect specifying acoustic correlates for the new features they propose.

From the large number of such discrepancies, particularly in the acoustic domain, we must conclude that the "natural" basis for phonology is largely ill-defined; though not necessarily undefinable.

E. Some Problems

The absence of precise definition of the empirical consequences of generative phonology precludes an effective recognition strategy; particularly for degraded speech signals. To fill this conceptual void, a number of motor theories and analysis-by-synthesis procedures have been forwarded. (See, Wathen-Dunn, 1967)

Illustration 1 removed due to nonreproducibility

However, I think it is fair to say none of these theories have:

(1) been shown comparable to known physiological, anatomical or neurological mechanisms in the manner required by the theories themselves. That is, they lack independent motivation.

(2) Seriously taken into account the generative power required for a perceptual strategy, given the known facts of synthesis.

(3) Separated recognition from identification functions adequate to account for the philosophical insight (See, Sayre, 1965)

(4) Approached precision and completeness.

(5) Avoided a surface taxonomic approach.

(6) Provided an internal, natural, evaluation measure as a principled way in refining alternative formulations.

(7) Worked, without ad hoc adjustment, in a manner commensurate with actual human speech perception.

Further, while some of the proposed theories reflect many of the known facts of language and its acquisition, neurophysiological and anatomical fact, and do not necessarily negate relevant speech performance factors, none as far as I have been able to ascertain, concurrently handle the following observations:

(1) Speech comprehension precedes speech production in order of acquisition.

(2) Aphasias can be almost exclusively receptive or expressive.

(3) Language is not learned normally by the congenitally deaf.

(4) Language is learned normally by the congenitally aphonic.

(5) Perceptual and articulatory strategies are unavoidably linearly dependent on time.

(6) Recognition vocabulary greatly exceeds production vocabulary.

(7) For some syntactic structures production is easy, but comprehension difficult, if not impossible. (e.g., triple embedding).

(8) For some phonological combinations or mechanical distortion of speech sounds, the corresponding articulation is extremely difficult or impossible; comprehension is normally simple. (e.g. compressed and expanded speech, tongue-twisters, etc.)

(9) We can recognize ill-formed utterances we produce, but we do not have to produce utterances to ascertain their well-formedness.

(10) Ambiguity is often recognized only after production.

(11) The same linguistically significant sound can be produced by different articulatory gestures. (e.g., /l/ is characteristically produced by some individuals by placing the apex of the tongue behind the lower, central incisors.)

This is not to say that a recognition strategy ought necessarily be concerned with a number of these observations. Rather, insofar as these and similar observations are correct, they will be compatible with a recognition strategy as it is interdependent with a psycholinguistic theory.

A recognition strategy is essential for not only psycholinguistic theory in general, but for ultimate justification of generative grammars. To make this claim is to blur the edges of the competence-performance methodological distinction. And, as Fano (1961) observes, "It turns out that the equipment required to generate efficient codes with long constraint spans is only moderately complex. The decoding equipment is inherently more complex."

F. A Proposed Solution

Nevertheless, I believe a solution is possible for the acoustic recognition problem. It does not necessarily follow that a solution will resolve either the general recognition problem, or provide a principled mechanism whereby an identity can be established between mental and physical entities, though successively refined approximations according to well-motivated principles must be provided for, at least by ostensive understanding of the physical requirements for mental constructs in a way that is deep and satisfying.

Based on general guidelines such as those suggested in this paper, I have been pursuing the acoustic recognition problem from a generative viewpoint.

This research, which was begun at the University of Minnesota, has led to the development of a theory of speech perception which is now entering laboratory verification stages. While several of the considerations included in this theory are too elaborate to discuss here, a sketch of the theory serves to illustrate considerations capable of satisfying most of the outstanding criticisms of contemporary perception strategies.

Work on acoustic formant transitions, largely conducted at the Haskins Laboratories, clearly indicates that the rate of change of formant transition is a significant variable in speech perception. Consequent analysis of the speech spectrum confirms that analysis of formants is most effectively accounted for in terms of differentiation of frequency with respect to time. That is, the generative power minimally required to account for transition slopes appears to be differential equations, with simple Fourier analysis being initially appealing (Flannagan, 1965). Initial applications of differentiation techniques (Rabiner, et. al. 1968, O'Heil, 1968, Oppenheim, 1970, Flannagan, 1965) have met with scant success.

Common to these approaches is the attempt to apply differentiation across the entire speech spectrum, using time as a base. However, studies in hearing (Stevens, 1961) indicate that below approximately 1,000 Hz certain aspects of the acoustic signal appear to be processed linearly, whereas above that frequency a logarithmic relationship seems to appertain. Moreover, wave amplitude seemed to be the only available index for the onset of temporal sequences.

For these and other reasons, I decided to pass the speech signal through a band-pass filter apparatus whereby the components above approximately 1,000 Hz could be separated from those below. To correspond to the requirements for both time and intensity information in initial recognition, the

signal above 1,000 Hz is differentiated by analogue procedures using intensity as a base. Below the 1,000 Hz mark, the signal is analogue differentiated by frequency with time as a base. Displays of the differentiated signals are available in parallel oscilloscopes. Then, the differentiated signals are passed through an analogue integrator, culminating in an oscillographic display which will, ideally, provide unique representations for input syllables according to their linguistic differences.

ILLUSTRATION 2 ABOUT HERE

Recognition procedures (differentiation stage) and identification procedures (integration and consequent linguistic stages) of this nature are capable of operating on compressed, masked, clipped, and somewhat expanded signals as well as signals with vastly different fundamentals. The output conditions of such procedures result in a syllabary. However, if the output signals vary in a principled way in accord with our linguistic experience, this provides not only a natural and independent basis for feature specification, but an indexing system for access to what could constitute a formidable syllabary.

Independent justification for speech perception procedures as outlined above can be found in compatability with neuroanatomical structures. I believe the recognition stage of this strategy is commensurate with the physical structure of the cochlea, where frequency/intensity differentiation is conducted in the scala vestibuli and ductus cochlearis; the closed endotic space and resiliance of Reissner's membrane serving to provide an averaging function necessary for an amplitude base. Usual place hypotheses are sufficient to account for the dual, frequency sensitivity requirement.

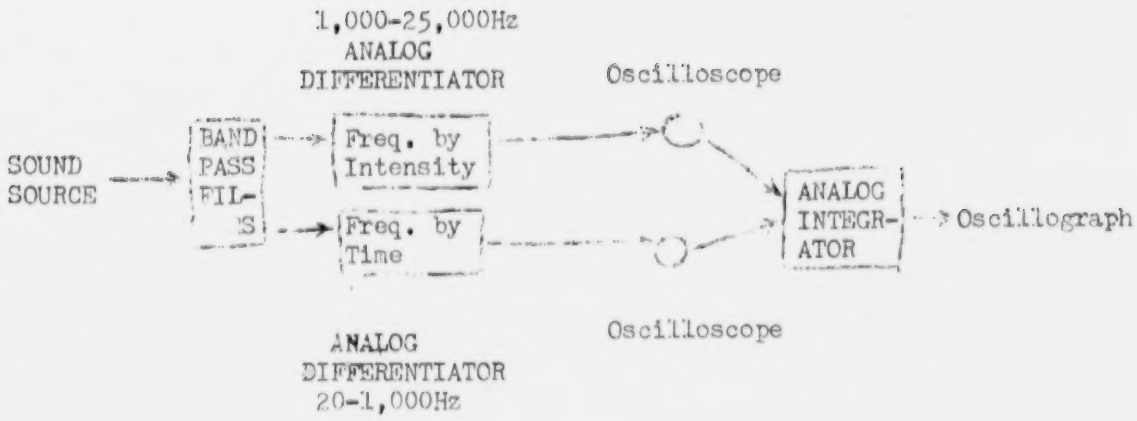

ILLUSTRATION TWO

A Schema of research instrumentation designed to investigate speech recognition. Parameters are adjustable. Ontological committment to functions is implied.

It should be noted, though, that the initial theory requires a closed, more viscous space in interaction with specific frequency sensitivity. Differentiation of frequency by time is postulated to occur as a result of activity associated with the scala tympani; suggesting interconnection of periotic spaces at the helicotrema and the relative thinness of the basilar membrane at its basal end to conduct amplitude transmissions immediately to the scala tympani and establish stable lower limits. Otherwise, serious syncronization problems could arise in simultaneous differentiation as the sound impulses pass around the apical extremety.

The integration stage of the hypothesized process is believed to occur more centrally and is subject, at this time, more to linguistic than neurophysiological comparison.

This "theory" is, to a large degree, speculation at this time. However, it is immediately subject to empirical test and can easily be disproven. Further work is now progressing and initial results ought to be available in the near future.

BIBLIOGRAPHY

Chomsky, Noam. Syntactic Structures. The Hague, Netherlands, 1957.

Chomsky, Noam. Current Issues in Linguistic Theory. The Hague, Netherlands, 1964.

Chomsky, Noam. "Formal Properties of Grammars," in R. Duncan, R. Bush and E. Galanter (Eds.) Handbook of Mathematical Psychology, Vol. II, pp. 323-418, New York, 1963.

Chomsky, Noam. Aspects of the Theory of Syntax. New York, 1965.

Chomsky, Noam. "Deep Structure, Surface Structure, and Semantic Interpretation," in D. Steinberg and L. Jakobovits (Eds.) Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology. New York, 1971.

Chomsky, Noam, and M. Halle. The Sound Pattern of English. New York, 1968.

DeLattre, P., A. M. Liberman, and F. Cooper. "Acoustic Loci and Transitional Cues for Consonants," Journal of the Acoustical Society of America, Vol. 27, No. 4, 769-773, July, 1955. Reprinted in I. Lehiste, Readings in Acoustic Phonetics. (below)

Fano, R., "Introductory Remarks," Proceedings of Symposia in Applied Mathematics, R. Jakobson (Ed.) Province, R.I., 1961, Vol. 12, 261-263.

Fillmore, C. "Types of Lexical Information," in D. Steinberg and L. Jakobovits (Eds.) Semantics... (below)

Flanagan, J., Speech Analysis, Synthesis and Perception, New York, 1965.

Fudge, E., "The Nature of Phonological Primes," Journal of Linguistics, Vol. 3, No. 1, April, 1967, 1-36.

Jakobson, R., C. Fant, and M. Halle. Preliminaries to Speech Analysis: The Distinctive Features and their Correlates. (revised edition), Cambridge, Mass., 1969.

Katz, J. and P. Postal. An Integrated Theory of Linguistic Descriptions. Cambridge, Mass., 1964.

Kisseberth, C. "A Global Rule in Klamath Phonology," Unpublished Ms. University of Illinois, Champaign.

Lakoff, G. "On Generative Semantics," in D. Steinberg and L. Jakobovits (Eds.) Semantics... (below)

Lakoff, G. "Presupposition and Relative Well-Formedness," in D. Steinberg and L. Jakobovits, (Eds.) Semantics... (below)

Lakoff, G. "The Arbitrary Basis of Transformational Grammar," in Language, Vol. 48, No. 1, March, 1972, 76-87.

BIBLIOGRAPHY cont'd

Langendoen, D. "Presupposition and Assertion in the Semantic Analysis of Nouns and Verbs in English," in D. Steinberg and L. Jakobovits, Semantics... (below)

Maclay, H. "Overview," in D. Steinberg and L. Jakobovits, Semantics... (below)

O'Neill, L. "The Representation of Continuous Speech with a Periodically Sampled Orthogonal Basis," IEEE Transactions on Audio and Electroacoustics, Vol. AU-17, No. 1, March, 1969, 14-21.

Oppenheim, A. "Speech Spectrograms Using the Fast Fourier Transform," IEEE Spectrum, August, 1970, 57-62.

Rabiner, L., H. Levitt, and A. Rosenberg, "Investigation of Stress Patterns for Speech Synthesis by Rule," Journal of the Acoustical Society of America Vol. 45, No. 1, 1969, pp. 92-101.

Sampson, G. "On the Need for a Phonological Base," Language, Vol. 46, No. 3, September, 1970, 586-625.

Sayre, K. Recognition: A Study in the Philosophy of Artificial Intelligence Notre Dame, 1965.

Steinberg, D. and L. Jakobovits. (Eds.) Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology. Cambridge, England, 1971.

Stevens, S. "The Psychophysics of Sensory Function," in W. Rosenblith (Ed.) Sensory Communication. Cambridge, Mass., 1959.

Turner, P. "Grammar for the Speaker and Hearer," Linguistics: An International Review, Vol. 57, May, 1970, 93-99.

Wathen-Dunn, W. (Ed.) Models for the Perception of Speech and Visual Form. Cambridge, Mass., 1964.

Whittaker, H. "On the Representation of Language in the Human Brain: Problems in the Neurology of Language and the Linguistic Analysis of Aphasia," UCLA Working Papers in Phonetics, 12., September, 1969.